Social Psychological and Personality Science
http://spp.sagepub.com/

Goals Can Be Threatened to Extinction : Using the Stroop Task to Clarify Working Memory Depletion Under Stereotype Threat

Keith A. Hutchison, Jessi L. Smith and Amber Ferris

Social Psychological and Personality Science 2013 4: 74 originally published online 27 March 2012
DOI: 10.1177/1948550612440734

The online version of this article can be found at:
http://spp.sagepub.com/content/4/1/74

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Society for Personality and Social Psychology
Association for Research in Personality
European Association of Social Psychology
Society of Experimental and Social Psychology

Social Psychological and
Personality Science
4(1) 74-81

DOI: 10.1177/1948550612440734
http://spps.sagepub.com

Goals Can Be Threatened to Extinction: Using the Stroop Task to Clarify Working Memory Depletion Under Stereotype Threat

Keith A. Hutchison[1], Jessi L. Smith[1], and Amber Ferris[1]

Abstract

Consensus is building that stereotype threat interferes with working memory, but how so? Grounded in the dual-process framework of Kane and Engle, the authors examined the extent to which stereotype threat interferes with one's ability to maintain task goals in working memory and one's ability to choose between conflicting responses. One hundred eighty-seven Montana State University (MSU) men were first given the Operation Span task (OSPAN) to assess working memory capacity, then engaged in the Stroop task under mostly incongruent or mostly congruent conditions. The Stroop task was presented as a measure of verbal processing skills (stereotype threat condition) or not (control condition). Stroop errors and reaction times were assessed. Results suggest that for people lower in working memory capacity, stereotype threat primarily interferes with internally maintaining task goals across trials. Implications for such stereotype threat-based distraction effects on performance in educational and workplace environments are discussed.

Keywords

stereotype threat, Stroop task, working memory

Science has well established that concern over confirming a stereotype leads to poor performance on stereotype-relevant tasks (Steele, Spencer, & Aronson, 2002). This negative effect of "stereotype threat" on performance is one of the more robust findings in social psychology (for a meta review, see Nguyen & Ryan, 2008). For example, under stereotype threat, African Americans perform worse on tests of intelligence (Steele & Aronson, 1995); women perform worse on tests of math (Spencer, Steele, & Quinn, 1999); men perform worse on tests of verbal skill (Seibt & Forster, 2004); and the elderly perform worse on memory tests (Levy, 1996). How stereotype threat brings about these deleterious effects on performance is less understood, as the situations leading to performance decrements can vary greatly (Smith, 2004).

The *working memory depletion* account for how stereotype threat affects performance has garnered the most attention in the literature (e.g., Schmader, 2010). Working memory is depleted under stereotype threat conditions (Schmader & Johns, 2003), such depletion mediates stereotype threat effects on performance (Beilock, Rydell, & McConnell, 2007) and people lower in working memory seem most vulnerable to stereotype threat effects (Regner et al., 2010). The goal of the current project was to understand just *how* stereotype threat "interferes" with mental resources. Thus, we hope to progress beyond simply stating that under stereotype threat someone has "divided attention" (Kaiser, Vick, & Major, 2006) or that working memory is "disrupted" (Bonnot & Croizet, 2007) and instead examine just how such working memory depletion manifests. Understanding just how working memory is affected by stereotype threat is important for scholars to better articulate the precise disadvantage that a stereotype threat context engenders for the stereotype target. Case in point, Engle and colleagues (Engle, 2002; Kane, Conway, Hambrick, & Engle, 2007) have argued that working memory capacity primarily reflects the ability to control attention, which serves to enhance appropriate information while simultaneously suppressing distraction. If stereotype threat interferes with executive attention, performance deficits may arise in any situation in which sustained attention and/or overcoming distraction is important. This involves not only the typical example of performance during standardized tests but also everyday situations for chronically threatened individuals including attending lectures, giving presentations, and even engaging in conversations (e.g.,

[1] Department of Psychology, Montana State University, Bozeman, MT, USA

Corresponding Author:
Keith A. Hutchison or Jessi L. Smith, Department of Psychology, Montana State University, P.O. Box 173440, Bozeman, MT 59717, USA
Email: khutch@montana.edu or jsismith@montana.edu

Holleran, Whitehead, Schmader, & Mehl 2010). Thus, gaining a nuanced understanding of stereotype threat processes can not only help explain why stereotype threat effects occur across different domains and why in some cases, only certain types of tests within a domain might show performance declines (e.g., Smith & White, 2002), but such knowledge can facilitate development of theoretically based interventions (e.g., Shapiro & Neuberg, 2007).

To accomplish these goals, we test *working memory depletion* using Kane and Engle's (2003) dual-process framework for performance in conflict tasks. According to Kane and Engle, performance on conflict tasks requires both goal maintenance and response competition processes. For instance, in the standard Stroop task (Stroop, 1935) in which people are told to quickly name aloud the color in which a word is written (e.g., the word *red* written in blue ink), participants must maintain the goal to selectively attend and respond to color while suppressing word information throughout the entire duration of the task. Failures in what is called "goal maintenance" over time can lead participants to accidentally default to the more habitual word naming response, resulting in errors on stimuli in which the word and color do not match (e.g., responding red in the above example). However, once such an incongruent (i.e., nonmatching) trial is encountered, this experience can serve as a goal reminder, allowing participants to be better prepared for the next trial. Therefore, frequent incongruent trials within a list can provide external goal support whereas the need for internal goal maintenance increases when most items in the list are congruent (e.g., the word *blue* written in blue ink). Consistent with this interpretation, Kane and Engle (2003) found that individuals low in working memory capacity (WMC) showed larger Stroop effects than high-WMC individuals, and that this difference was especially pronounced under mostly congruent conditions. Similarly, Hutchison (2011) found that low-WMC individuals showed larger Stroop effects in mostly congruent than in mostly incongruent lists whereas high-WMC individuals were not influenced by list context, suggesting that low-WMC individuals are more dependent upon external support for sufficient goal maintenance. Moreover, low-WMC individuals are more prone to mind wandering (McVay & Kane, 2009) and have difficulty maintaining task goals in the presence of interference from distracting information (Kane et al., 2007).

Even if an individual actively maintains the task goal, one still must overcome the competing responses present on incongruent trials. In the Stroop task, resolving such "response competition" requires both inhibiting the inappropriate word response and selecting the appropriate color response. Kane and Engle (2003) argued that individuals low in WMC require more time to inhibit the inappropriate word response, which leads to prolonged reaction times (RT) to incongruent stimuli. Thus, even under mostly incongruent conditions in which goal maintenance is externally reinforced, WMC differences may still arise in RT due to deficient ability to resolve response competition among low-WMC individuals.

To the extent that stereotype threat reduces working memory, we aim to clarify just how this "reduction" manifests. To do this, we draw from the dual-process framework described above and employ the Stroop task. A handful of recent stereotype threat studies have employed variations of the Stroop task (Carr & Steele, 2010; Inzlicht & Kang, 2010; Inzlicht, McKay, & Aronson, 2006). Both Carr and Steele and Inzlicht and Kang administered the Stroop task to women following a math task and found that women under stereotype threat had a larger Stroop effect; however, because both studies examined only RTs and did not vary list congruency, it was not possible to distinguish between impaired goal maintenance and response competition mechanisms. Similarly, Inzlicht, McKay, and Aronson (2006) found that African American participants under stereotype threat also showed elevated Stroop RTs. However, as with the other studies, Inzlicht et al. did not examine error rates across conditions nor manipulate list congruency. Thus, in the few cases in which Stroop effects were examined within a stereotype threat paradigm, the conclusions that can be drawn are limited because list congruency was not manipulated and the full range of effects was not examined.

Examining the direct influence of stereotype threat on Stroop performance can also produce a marked advantage over past studies. The few studies to use Stroop in a stereotype threat study have administered the Stroop task either before (Inzlicht et al., 2006) or after (Carr & Steele, 2010; Inzlicht & Kang, 2010) the stereotype-related task, making it unknown whether stereotype threat is operating during the Stroop test itself. Instead, such deficits could reflect either anticipation of an upcoming threatening situation or rumination over impaired performance during an earlier threatening situation. Neither of these is equivalent to stereotype threat experience *during* the test itself. Using the Stroop task as the stereotype-relevant performance test, we know stereotype threat is operating during the Stroop task itself.

Another advantage of examining stereotype threat while performing the Stroop task is that it is a domain-general attention task that has been shown to correlate with performance across a wide range of tasks. Stroop effects predict performance on tasks that require source monitoring in memory (Sommers & Huff, 2003), visual attention (Hutchison, 2007), working memory (Hutchison, 2011; Kane & Engle, 2003), and spatial manipulation (Hutchison, Balota, & Duchek, 2010). For instance, Hutchison, Balota, and Duchek (2010) found that the percentage of incongruent Stroop errors committed by their sample of healthy older adults and individuals diagnosed with Dementia of the Alzheimer's Type significantly correlated with 17 out of their 18 other measures of attention and memory performance. They argued that Stroop performance likely captures a fundamental attentional control ability which underlies performance across a wide range of cognitive tasks. Thus, using the Stroop task as the performance test itself, findings should generalize across many stereotype domains such as math, verbal, memory, intelligence, and engineering.

Perhaps, the most important reason to investigate stereotype threat within the Stroop task is that this task allows us to examine specific hypotheses regarding the presumed nature of working memory disruption under threat. Specifically, a *distraction* hypothesis would predict that stereotype threatened individuals become distracted and start to mind wander during the task, which interferes with their ability to maintain the task goal of naming colors not words. If stereotype threat causes distraction, threat effects should emerge specifically on error rates within mostly congruent lists in which task goals must be internally supported. Such a pattern of results would suggest stereotype threat impairs the target individuals' ability to stay focused on their test-taking strategy and goals for successful test performance. In contrast, a *response competition* hypothesis would predict that stereotype threat impairs peoples' ability to resolve the competition created by the incorrect, but habitual, word response. This response competition hypothesis predicts stereotype threat effects emerge in Stroop RTs within mostly incongruent lists in which goals are externally supported. Such a pattern of results would suggest that stereotype threat impairs the speed with which target individuals can override their (incorrect) habitual responses. Finally, stereotype threat may impair both the ability to avoid distraction and to resolve response competition, in which both hypotheses would be correct and stereotype threat effects should emerge both in Stroop error rates under mostly congruent lists and in Stroop RTs within mostly incongruent lists. In addition to these possible predictions, we aim to determine the role of trait levels of WMC. Some research has shown that those high in WMC have enough resources to be buffered from stereotype threat effects (see Regner et al., 2010 for a recent example). Alternatively, there is research showing that those high in WMC are likely to choke when under performance pressure (Beilock & Carr, 2005). Thus, it is unclear whether stereotype threat should cause high-WMC individuals to perform similar to nonthreatened low-WMC individuals (i.e., choking) or if high-WMC individuals would instead be immune to stereotype threat effects that greatly impair low-WMC individuals. According to the distraction hypothesis, stereotype threat leads to distraction which causes people to accidently default to the habitual process of word reading, rather than color naming. This should produce an increase in errors to incongruent stimuli. As such, we expect that low-WMC individuals should be more vulnerable to stereotype threat-induced distraction compared to high-WMC individuals, since low-WMC individuals are already prone to distraction (McVay & Kane, 2009). If so, a majority of the effect should be due to a decrease in the already poorer performance of low-WMC individuals. Certainly, it is possible that stereotype threat could also lead to high-WMC individuals being distracted. However, it is unlikely that they would show equal or more threat-induced distraction than low-WMC individuals. The same argument applies to the response-competition hypothesis, in which low-WMC individuals are suggested to have deficits overcoming habitual responses even when goals are externally maintained (Kane & Engle, 2003). Thus, we predict greater stereotype threat deficits among low-WMC individuals.

To test these predictions, we utilized the stereotype that men are less verbally skilled than women (e.g., Guimond & Roussel, 2001; Hyde & Kling, 2001; Hyde & Linn, 1988) following Seibt and Forster's (2004) procedure for examining men's performance on verbal tests following gender stereotype threat. Pretesting indeed showed the Stroop color naming test was believably a ''verbal processing'' test. In this way, we could ensure that stereotype threat was operating during Stroop performance, rather than assessing Stroop effects before or after completion of the stereotype-domain relevant test, as has been done in previous studies (Carr & Steele, 2010; Inzlicht & Kang, 2010).

Method

Participants and Procedure

A total of 187 men ($M = 21.2$ years old, 88.5% Caucasian) participated in exchange for course credit. Upon arrival, participants were first given the automated version of the operation span WMC task (Unsworth, Heitz, Schrock, & Engle, 2005). In this task, participants answer math problems (e.g., $2 \times 4 + 1 = 9$) as quickly as possible. After each response, they are presented with a letter for 800 ms to hold in memory for a later test. After three to seven problems, participants are presented with a 3×4 matrix of letters and asked to click on the presented letters in the order in which they were shown. Set sizes varied from three to seven problems and participants received three sets of each size, with set size order randomized for each participant. An individual's operation span score is the sum of all perfectly recalled sets. For example, if a person correctly recalled all the letters from 3/3 three-item sets, from 2/3 four-item sets, and from 1/3 five-item sets, their span score would be 22 (3 + 3+ 3+ 4 +4 + 5). Unsworth et al. (2005) demonstrated that this version of OSPAN correlates well with other measures of WMC and has both good internal consistency ($\alpha = .78$) and test–retest reliability (.83).

Next, participants were randomly assigned to one of the four conditions in a 2 (stereotype threat vs. no threat) $\times$ 2 (mostly congruent list vs. mostly incongruent list) between-subjects design with 94 participants in the mostly congruent condition ($n = 47$ control; $n = 47$ stereotype threat) and 92 participants in the mostly incongruent condition ($n = 46$ control; $n = 46$ stereotype threat).

Modeled after Seibt and Forster (2004), men in the stereotype threat condition were told they were about to engage in a test measuring ''verbal skills of men and women'' and marked their gender before the test. Men in the control condition were told that the test measured ''processing skills'' and were not asked to mark their gender until after the test.

All participants then took the computer-based color-naming Stroop task. The Stroop stimuli were taken from Spieler, Balota, and Faust (1996) and consisted of both color words (red, green, blue, and yellow) and neutral words (bad, deep, poor, and legal) matched to the color words in onset phoneme characteristics and printed word frequency (Kucera & Francis, 1967). After 16 randomly presented practice trials, participants received 200 experimental trials consisting of 24 incongruent trials (e.g., the word *red* written in blue), 24 congruent trials (e.g., the word *blue* written in blue), 32 neutral trials (e.g., the word *deep* written in blue), and 120 filler trials. Following Hutchison (2011), the filler trials were not analyzed and consisted of the same four color words presented solely in their congruent or incongruent colors in order to create mostly congruent and mostly incongruent lists. The mostly congruent list consisted of 144 congruent stimuli and 24 incongruent stimuli whereas the mostly incongruent list consisted of 24 congruent stimuli and 144 incongruent stimuli. Stroop stimuli were presented using E-prime software (Schneider, Eschman, & Zuccolotto, 2002).

Participants were instructed to name the ink color, but not the word itself and to respond as quickly and accurately as possible. Trials were presented in a fixed random order. Targets were presented in the center of the computer screen for 4,000 ms or until a response and were separated by a 1,500 ms intertrial interval. RTs from the onset of the target stimulus until the onset of the vocal response were coded by a model 300 Psychology Software Tools (PST) serial response box. Following the target response, response errors were coded by an experimenter who sat next to the participant, held a sheet with the correct (color) responses for each trial and coded participants' responses as (a) correct response, (b) response error, or (c) microphone error. Holding the correct answers in front of them allowed the experimenters to accurately record participant's responses. The experimenter coded the trial as a microphone error if the microphone failed to pick up the vocal response and a response error if the participant either said the wrong color entirely or produced a blend (e.g., *green–blue*).

Results

Only correct responses were considered for the RT analyses. A separate mean and standard deviation were computed for congruent, neutral, and incongruent trials for each participant. The nonrecursive outlier removal procedure, suggested by Van Selst and Jolicoeur (1994), was used in which the criterion for outlier removal is adjusted based upon the number of correct observations in each condition for each participant to correct for the bias that exists with more traditional outlier removal procedures (see Van Selst & Jolicoeur, 1994, for more discussion). This procedure removed 2.5% of the RTs for correct trials.

For participants in the control condition, operation span scores were as follows: in the mostly congruent list: $M = 38.1$, $SE = 2.1$; mostly incongruent list: $M = 40.9$, $SE = 2.7$. For the stereotype threat condition, operation span scores were as follows: mostly congruent list: $M = 39.6$, $SE = 2.5$ and mostly incongruent list: $M = 38.8$, $SE = 2.5$. None of the operation span scores differed across the four experimental conditions (all t values < 1).

Basic regression models examining Stroop effects (i.e., incongruent–congruent conditions) in both errors and RTs were created with the main effect of stereotype condition (coded as *stereotype threat* = 1; *no stereotype threat* = −1), the main effect of list congruency (coded as *mostly congruent* = 1; *mostly incongruent* = −1), main effect of OSPAN WMC (centered and measured continuously); the three two-way interaction terms and the one three-way interaction term. Three participants with microphone errors greater than 25% and two participants with incomplete data were dropped from analyses, resulting in 182 total participants. RTs and error rates to neutral items did not vary as a function of threat, regardless of list (all t values < 1.1) so neutral items are not discussed further.

Testing the Distraction Hypothesis

To test the distraction hypothesis, Stroop effects in errors were regressed on the model. The overall model was significant $R^2 = .65$, $F(7, 174) = 18.41$, $p < .001$, and revealed the typical effects of list congruency and WMC on Stroop errors, whereby participants' Stroop effects were greater in the mostly congruent list than the mostly incongruent list ($\beta = -.54$, $t = 9.33$, $p < .01$), and participants lower in WMC had larger Stroop effects than those higher in WMC ($\beta = -.23$, $t = -3.87$, $p < .01$). Also, replicating Hutchison (2011), these two main effects were qualified by their two-way interaction ($\beta = -.22$, $t = -3.81$, $p < .0$). Stratified analyses by list-type showed WMC differences in Stroop effects were pronounced among those receiving the mostly congruent list ($\beta = -.39$, $t = -4.16$, $p < .01$) but absent among those receiving the mostly incongruent list ($\beta = -.01$, $t = -0.06$, $p = .95$). As well, there was a main effect of stereotype threat ($\beta = .12$, $t = 2.11$, $p < .05$) such that men in the stereotype threat condition had a larger Stroop effect than those in the control condition. This stereotype threat main effect was qualified by an interaction with list condition ($\beta = .13$, $t = 2.29$, $p < .05$). Again, stratified analyses by list-type showed stereotype threat increased Stroop errors among those in the mostly congruent list ($\beta = .24$, $t = 2.55$, $p < .05$), but not among those in the mostly incongruent list ($\beta = -.02$, $t = -0.17$, $p = .86$). There was also a marginal two-way interaction between threat condition and WMC ($\beta = -.11$, $t = -1.93$, $p = .054$).

Importantly, however, these main effects and two-way interactions were qualified by a significant three-way interaction between stereotype threat, list congruency, and WMC ($\beta = -.12$, $t = -1.99$, $p < .05$). Predicted values for this three-way interaction are plotted in Figure 1. To decompose the interaction, we tested the Simple Stereotype Threat × WMC interaction separately for each list and then, if significant, determined the effects of stereotype threat separately for those relatively high or low in WMC (using 1 *SD* above and below 39.50). For the mostly congruent list, the two-way Stereotype Threat × WMC interaction was significant ($\beta = -.20$, $t = -2.14$, $p < .05$). Following guidelines by Aiken and West (1991) , comparison of the simple slopes in the mostly congruent list at 1 *SD below* the mean WMC revealed a significant slope for the stereotype threat condition, ($\beta = -.30$, $t = -2.15$, $p < .05$) with Stroop effects increasing under stereotype threat. However, the simple slope analysis at 1 *SD above* the mean WMC did not reveal significant difference ($p = .86$, see Figure 1). Thus, within the mostly congruent list, in which internal goal maintenance is required for successful

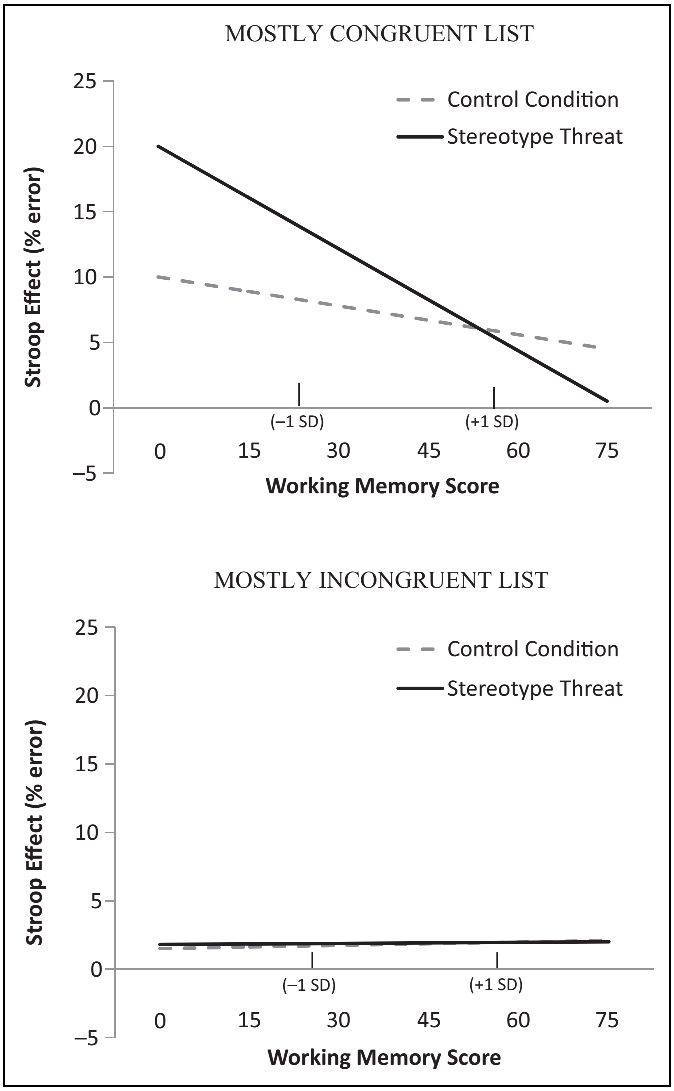

Figure 1. Predicted values of Stroop effects in error rates by working memory capacity (WMC) for stereotype threat and control participants within mostly congruent (top) or mostly incongruent (bottom) lists.

performance, those with lower WMC made more Stroop errors in the stereotype threat condition compared to the control (no threat) condition. In contrast to the mostly congruent list, the two-way Stereotype Threat × WMC interaction was not significant among those in the mostly incongruent list ($\beta = .01$, $t = 0.59$, $p = .95$), suggesting no effect of stereotype threat on Stroop performance (see Figure 1).

As predicted by the distraction hypothesis, these results suggest that stereotype threat interferes with goal maintenance particularly among low-WMC individuals, presumably due to their increased susceptibility to distraction (McVay & Kane, 2009). Specifically, low-WMC individuals showed increased Stroop errors under stereotype threat for the mostly congruent list. Within the mostly incongruent list (see bottom of Figure 1), stereotype threat had no effect on Stroop performance, regardless of WMC.

Testing the Response Competition Hypothesis

To test the response competition hypothesis, Stroop RT effects were regressed on the model, which was also significant $R^2 = .53$, $F(7, 174) = 28.95$, $p < .001$. However, the only significant effects were a main effect of WMC ($\beta = -.11$, $t = -2.01$, $p < .05$) in which Stroop RT effects were larger for low-WMC individuals and a main effect of list ($\beta = .72$, $t = 13.949$, $p < .01$) in which Stroop effects were larger in the mostly congruent list. These findings replicate previous research (e.g., Hutchison, 2011). No other main effects or two-way interactions emerged (all t values < 1), and the three-way interaction did not approach statistical significance ($p = .29$).

Discussion

As predicted by the distraction hypothesis, stereotype threat produced a significant increase in Stroop errors among those low in WMC only under conditions in which the task goal had to be internally maintained (i.e., mostly congruent list). In contrast, when the task goal was externally supported through frequent exposure to incongruent stimuli, no such disruption occurred. These findings support the hypothesis that stereotype threat impairs performance on conflict tasks by disrupting peoples' ability to maintain focus throughout the task. The distraction caused by stereotype threat–related thoughts leads to a loss of the appropriate goal to suppress habitual, yet incorrect, response tendencies.

Our results did not support the response competition hypothesis, neither low- nor high-WMC individuals showed stereotype threat-induced response competition effects. Therefore, we found evidence that stereotype threat exerts its debilitating effect primarily through goal neglect, but found no evidence that stereotype threat weakens one's ability to resolve response competition. Certainly, more research is needed to replicate our results for the distraction and response competition hypotheses. It is possible, for example, that with other stereotypes (e.g., race-related competence stereotypes) or among people who are more commonly targets of stereotypes (e.g., African Americans) response competition deficits would also emerge. It is also possible that the manner in which stereotype threat is activated may result in more or less impairments because of distraction and/or response competition. According to Stone and McWhinnie (2008), a subtle (or implicit) stereotype threat impairs working memory because the person is focused on detecting bias. In contrast, when the threatening stereotype is activated in a more blatant, explicit manner (e.g., by reminding people that gender stereotypes exist), the result is worry and performance-avoidance goal adoption (e.g., Smith, Sansone, & White, 2007). In the current project, we examined our hypotheses with a majority group who are less commonly victims of stereotype threat (but see Koenig & Egly, 2005; Smith & White, 2002; Stone, Lynch, Sjomeling, & Darley, 1999) using a subtle manipulation to activate the stereotype, suggesting that even with relatively less "threat in the air" the distraction hypothesis holds true.

Our results fit well with research on mind wandering. Individuals low in WMC are more prone to mind wandering (McVay & Kane, 2009) and such mind wandering predicts intrusion errors within a vigilance task. Moreover, mind wandering seems to play an important role in the stereotype threat–performance relationship (Mrazek et al., 2011). Integrating this line of work with our own suggests that the minds of low-WMC individuals under stereotype threat begin to wander during the task, causing goal neglect failures which, in turn, cause intrusions of inappropriate habitual word reading responses over the more appropriate color-naming responses. The content of such mind wandering is still unknown, but may consist of metacognitions regarding performance (Schmader, Forbes, Zhang, & Mendes 2009), bias detection (Kaiser et al., 2006), interpreting internal affective responses (Schwarz & Clore, 1983), and ruminations of past stereotypic experiences.

We began this work as a straightforward test of the working memory depletion account of stereotype threat effects. In fact, the working memory depletion account informed by Kane and Engle's (2003) dual-process theory accurately predicted that stereotype threat would produce increased errors only under mostly congruent lists and that these effects would occur exclusively for low-WMC individuals. However, an alternative "mere effort" account (Jamieson & Harkins, 2007) may also explain an increase in Stroop errors under threat. According to mere effort, people under stereotype threat have increased anxiety over performance evaluation, which leads to a potentiation of prepotent responses and increased effort to correct initial mistakes. Although mere effort theory accurately predicts more errors overall when under stereotype threat, it is unclear why the increase in errors would only occur for low-WMC individuals who receive mostly congruent lists (as we found here). Modifications to mere effort theory would therefore be needed to fully explain the current data, which will be a fruitful area for future research.

In summary, results from the current study clearly demonstrate that under stereotype threat, a disruption in goal maintenance occurs during task engagement, especially among those low in WMC. We posit that subtly triggered stereotype threat effects reflect goal-neglect problems in which individuals (particularly those low in WMC) become distracted following the threat and begin mind wandering. This mind wandering then causes goal-neglect, which impairs performance in attention-demanding situations such as those involving planning, troubleshooting, technical difficulty, or novel sequences of action (Norman & Shallice, 1986).

Understanding that stereotype threat depletes the target's ability to maintain task goals has implications for understanding why stereotype threat influences such a wide range of outcomes. Stereotype threat–induced distractions are everywhere, for example, in television commercials (Davies, Spencer, & Steele, 2002) in situations where a person is in the numerical minority (Roberson, Deitch, Brief, & Block, 2003) or is the target of a hostile joke (Oswald & Harvey, 2001). Any task in which focal attention is necessary is therefore prone to distraction-based stereotype threat effects. Such tasks include studying for an exam, preparing for or giving a presentation, listening attentively during a meeting, problem solving, taking a driving test, and fielding questions during one's dissertation defense. Even in relatively structured settings such as taking a 50-min subsection of a graduate record examination test, there are ample opportunities for fluctuations in attentional focus. Indeed, McVay and Kane (2011) recently demonstrated that mind wandering mediated the influence of WMC on reading comprehension, suggesting that a major factor in reading comprehension is the ability to exert control over potentially intruding thoughts.

Our data suggest mind wandering is an important reason why stereotype threat produces such varied outcomes as counterproductive leadership communication styles (von Hippel, Wiryakusuma, Bowden, & Shochet, 2011) and impaired driving skills among women (Yeung & von Hippel, 2008). Indeed, we suggest that stereotype threat effects will be greatest not only among those chronically prone to distraction (i.e., individuals low in WMC or diagnosed with attention deficit hyperactivity disorder (ADHD)), but to all individuals during states of impaired WMC such as that during divided attention, ego-depletion (Carr & Steele, 2010), or off-peak circadian rhythms (Hasher, Zacks, & May, 1999).

While it may be impossible to nullify a stereotype in a given culture, it is possible to arm employers and educators with training and information on how to optimize a given performance setting for stereotype targets (e.g., by minimizing other distractions while working; by placing goal reminder cues in the environment). In addition, education regarding how stereotype threat enhances distraction may empower stereotype targets to take measures to avoid such distraction during performance, for example, by catching themselves mind wandering and refocusing on the task. Future studies should examine whether such foreknowledge can indeed reduce decrements caused by stereotype threat. Overall, our work underscores the importance of setting up supportive environments that will reinforce task goals to prevent them from being (stereotype) threatened to extinction.

Acknowledgments

The authors wish to thank Katie Vander Vos, Carli Dickson, and the Motivation and Diversity Lab for their assistance in collecting data used in this study.

Authors' Note

The manuscript contents are solely the responsibility of the authors and do not necessarily represent the official view of NSF.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Financial Disclosure/Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article:

Authorship of this article was made possible in part by Grant Number HRD-1036767 from the National Science Foundation awarded to the second author.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: SAGE Publications.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, *16*, 101–105.

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviations, and spillover. *Journal of Experimental Psychology: General*, *136*, 256–276.

Bonnot, V., & Croizet, J. C. (2007). Stereotype internalization and women's math performance: The role of interference in working memory. *Journal of Experimental Social Psychology*, *43*, 857–866. doi: 10.1016/j.jesp.2006.10.006

Carr, P. B., & Steele, C. M. (2010). Stereotype threat affects financial decision making. *Psychological Science*, *21*, 1411–1426.

Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology*, *88*, 276–287.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Guimond, S., & Roussel, L. (2001). Bragging about one's school grades: Gender stereotyping and students' perception of their abilities in science, mathematics, and language. *Social Psychology of Education*, *4*, 275–293.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and cognition. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). Cambridge: MIT Press.

Holleran, S. E., Whitehead, J., Schmader, T., & Mehl, M. R. (2010). Talking shop and shooting the breeze: A study of workplace conversation and job disengagement among STEM faculty. *Social Psychological & Personality Science*, *2*, 65–71.

Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 645–662. doi: 10.1037/0278-7393.33.4.645

Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 851–860. doi:10.1037/a0023437

Hutchison, K. A., Balota, D. A., & Duchek, J. M. (2010). The utility of Stroop task switching as a marker for early stage Alzheimer's Disease. *Psychology and Aging*, *25*, 545–559.

Hyde, J. S., & Kling, K. C. (2001). Women, motivation, and achievement. *Psychology of Women Quarterly. Special Issue: Women's Lives, Feminist Research, and the Decade of Behavior*, *25*, 364–378. doi:10.1111/1471-6402.00035

Hyde, J. S., & Linn, M. C. (1988). Gender difference in verbal ability: A meta-analysis. *Psychological Bulletin*, *104*, 53–69. doi:10.1037/0033-2909.104.1.53

Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with threats to social identity affects aggression, eating, decision making, and attention. *Journal of Personality and Social Psychology*, *99*, 467–481. doi:10.1037/a0018951

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science*, *17*, 262–269. doi:10.1111/j.1467-9280.2006.01695

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology*, *93*, 544–564.

Kaiser, C. R., Vick, S. B., & Major, B. (2006). Prejudice expectations moderate preconscious attention to cues that are threatening to social identity. *Psychological Science*, *17*, 332–338.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–48). Oxford, England: Oxford University Press.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47–70. doi:10.1037/0096-3445.132.1.47

Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles*, *52*, 489–496.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Levy, B. (1996). Improving memory in old age through implicit self stereotyping. *Journal of Personality and Social Psychology*, *71*, 1092–1107.

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-sontrol task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 196–204.

McVay, J. C., & Kane, M. J. (2011). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*. Advance online publication. doi: 10.1037/a0025250

Mrazek, M. D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology*, *47*, 1243–1248. doi: 10.1016/j.jesp.2011.05.011

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. *4*, pp. 1–18). New York, NY: Plenum.

Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*, 1314–1334.

Oswald, D. L., & Harvey, R. D. (2000–2001). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology: Developmental, Learning, Personality, Social*, *19*, 338–356.

Regner, I., Smeding, A., Gimmig, D., Thinus-Blanc, C., Monteil, J.-M., & Huguet, P. (2010). Indivdiual differences in working memory moderate stereotype threat effects. *Psychological Science*, *21*, 1646–1648.

Roberson, L., Deitch, E. A., Brief, A. P., & Block, C. J. (2003). Stereotype threat and feedback seeking in the workplace. *Journal of Vocational Behavior*, *62*, 176–188.

Schmader, T. (2010). Stereotype threat deconstructed. *Current Directions in Psychological Science*, *19*, 14–18.

Schmader, T., Forbes, C. E., Zhang, S., & Mendes, W. B. (2009). A meta-cognitive perspective on cognitive deficits experienced in intellectually threatening environments. *Personality and Social Psychology Bulletin*, *35*, 584–596.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*, 440–452.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Seibt, B., & Forster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, *87*, 38–56.

Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review, 11*, 107-130. doi: 10.1177/1088868306294790

Smith, J. L. (2004). Understanding the process of stereotype threat: A review of mediational variables and new performance directions. *Educational Psychology Review*, *16*, 177–206.

Smith, J. L., Sansone, C., & White, P. H. (2007). The stereotyped task engagement process: The role of interest and achievement motivation. *Journal of Educational Psychology*, *99*, 99–114.

Smith, J. L., & White, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles*, *47*, 179–191.

Sommers, M. S., & Huff, L. M. (2003). The effects of age and dementia of the Alzheimer's type on phonological false memories. *Psychology and Aging*, *18*, 791–806.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4–28.

Spieler, D. H., Balota, D. A., & Faust, M. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 461–479.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of Afircan Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. *34*, pp.379–440). San Diego, CA: Academic Press.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology*, *77*, 1213–1227.

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology*, *44*, 445–452.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. doi: 10.1037/h0054651

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 498–505.

Van Selst, M., & Jolicoeur. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology*, *47A*, 631–650.

Von Hippel, C., Wiryakusuma, C., Bowden, J., & Shochet, M. (2011). Stereotype threat and female communication styles. *Personality and Social Psychology Bulletin*, *37*, 1312–1324. doi:10.1177/0146167211410439

Yeung, N. C. J., & von Hippel, C. (2008). Stereotype threat increases the likelihood that female drivers in a simulator run over jaywalkers. *Accident Analysis & Prevention*, *40*, 667–674.

Bios

Keith A. Hutchison is an associate professor of psychology who studies attention, memory, and language in young adults, older adults, and individuals diagnosed with dementia of the Alzheimer's type. He completed his PhD at SUNY Albany and a 2-year postdoctorate position at Washington University in St. Louis.

Jessi L. Smith is an associate professor of psychology who studies gender, stereotype threat, self-regulation of motivation, and interpersonal and achievement goals. She completed her PhD at the University of Utah.

Amber Ferris is currently enrolled in the Masters in Psychological Science program at Montana State University. Her primary mentor is Jessi Smith.